

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 April 2003 (10.04.2003)

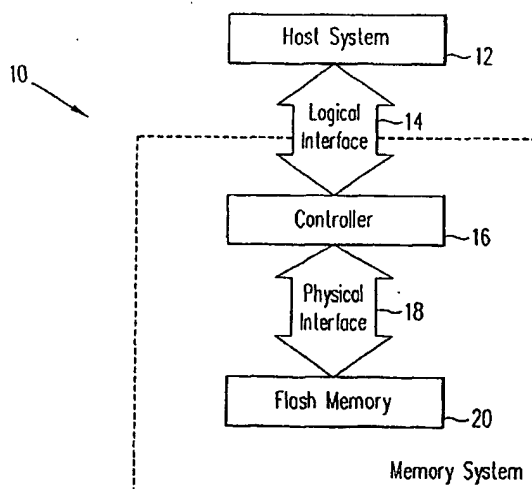
PCT

(10) International Publication Number  
**WO 03/029951 A2**

- (51) International Patent Classification<sup>7</sup>: **G06F 3/06** (74) Agents: **BERESFORD, Keith, Denis, Lewis et al.**; Beresford & Co., 2-5 Warwick Court, High Holborn, London WC1R 5DH (GB).
- (21) International Application Number: **PCT/GB02/04393**
- (22) International Filing Date:  
27 September 2002 (27.09.2002) (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
0123416.0 28 September 2001 (28.09.2001) GB
- (71) Applicant (*for all designated States except US*): **LEXAR MEDIA, INC.** [US/US]; 47421 Bayside Parkway, Fremont, CA 94538 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (*for US only*): **GOROBETS, Sergey, Anatolievich** [GB/GB]; 1F1 16 East Mayfield, Edinburgh EH9 1SE (GB).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: **NON-VOLATILE MEMORY CONTROL**



(57) Abstract: According to an embodiment of the present invention, there is provided a method and apparatus for use in a memory system having a non-volatile memory and a controller for limiting the number of non-volatile memory arrays from a plurality of available arrays accessed at one time, wherein the method comprises implementing a pipelining sequence for transferring data to and from the non-volatile memory arrays and limiting the number of active arrays operating at one time, the arrangement being such that the controller waits for the at least one of the arrays to complete before initiating the transfer to and from a further array.



**Published:**

— without international search report and to be republished  
upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

NON-VOLATILE MEMORY CONTROL

The present invention relates to a solid state memory system for data storage and retrieval, and to a memory controller for controlling access to a non-volatile memory of a solid state memory system. In particular, the present invention relates to a method of fast access of the data in the memory system with a precise control of power consumption, and including the control of flash memory accesses.

It is known to use solid state memory systems to try to emulate magnetic disc storage devices in computer systems. It is an aim of the industry to try to increase the speed of operation of solid state memory systems to better emulate magnetic disc storage.

A typical memory system comprises a non-volatile (Flash) memory and a controller. The memory has individually addressable sectors where a memory sector is a group of flash memory locations which is allocated for storage of one Logical Sector. A memory sector need not be a physical partition within Flash memory, nor contiguous Flash memory locations, so that the memory sector address may be a virtual address conveniently used by the controller. The controller writes data structures to and reads data structures from the memory, and

translates logical addresses received from the host to physical (virtual) addresses of the memory sectors in the memory.

An example of such a memory system is illustrated by the Memory System of WO 00/49488. In Figure 1 (prior art) there is illustrated the timing of various operations involved in a multiple sector write to interleaved flash chips forming a flash array described for the memory system of WO 00/49488.

However in many systems, and in particular systems such as portable computers, the maximum level of electrical current is a very important parameter defining the system design, efficiency and cost. For systems which include memory storage devices the number of flash memory chips active at the time is a major factor defining the current level. It is therefore important to control the maximum value of electrical current level to avoid high peaks which can cause higher requirements to the host system power supply. It is also important to be able to change the maximum current level and to compromise on performance if required.

An object of the invention is to provide a method of controlling flash memory accesses in a memory system configured to use concurrent operation in different flash memory arrays, and allowing performance to be changed

easily in systems supporting concurrent flash operations in different Flash arrays.

According to the present invention there is provided a method for use in a memory system having a non-volatile memory and a controller for limiting the number of non-volatile memory arrays from a plurality of available arrays accessed at one time, wherein the method comprises implementing a pipelining sequence for transferring data to and from the non-volatile memory arrays and limiting the number of active arrays operating at one time, the arrangement being such that the controller waits for at least one of the arrays to complete before initiating the transfer to and from a further array.

The present invention also provides a memory system having a non-volatile memory incorporating a plurality of memory arrays and a controller which is arranged to implement the method of the immediately preceding paragraph.

These and other aspects of the invention will become apparent from the following description taken in combination with the accompanying drawings in which:  
Figure 1 shows a schematic representation of timing :  
operations of a prior art memory system;

Figure 2 illustrates a memory system and associated host system in which the present invention is implemented;

Figure 3 shows a schematic representation of the hardware architecture of the controller of the memory system of Figure 2;

Figure 4 shows a schematic representation of the firmware executed in the microprocessor of the controller of Figure 3;

Figure 5A shows a schematic representation of the data write operation used by the controller of Figure 3;

Figure 5B illustrates read and write pointer operations;

Figure 6 is a schematic representation of the hierarchy of mapping structures of the address translation process used by the controller of Figure 3;

Figure 7 is a schematic representation of a method of scheduling the transfer of sector data according to a first embodiment of the present invention;

Figure 8 is a schematic representation of a method of scheduling the transfer of sector data according to second embodiment of the present invention;

Figure 9 is a schematic representation of a method of scheduling the transfer of sector data according to third embodiment of the present invention;

Figure 10 shows a schematic representation of a method of scheduling the transfer of sector data according to fourth embodiment of the present invention;

Figure 11 is a schematic representation of a method of scheduling the transfer of data whilst limiting the number of active arrays according to the present invention;

Figure 12 shows an alternative memory system arrangement in which the present invention may be implemented.

A Flash disk device, such as that shown in Figure 2, is a memory system which presents the logical characteristics of a disk storage device to a host system 12, and which uses Flash semiconductor memory 20 as its physical data storage medium. A Flash disk memory system 10 requires a controller 12 to manage the physical storage medium of the system 10 according to algorithms which create the logical characteristics of a disk and, in this case, it is the flash memory 20 and controller 16 which are connected by physical interface 16 which form

the memory system 10. The controller 16 of the memory system 10 connects the system 10 to the host 12 via logical interface 14.

In this case the flash memory 20 comprises a plurality of flash chips which are formed of a plurality of flash blocks. The logical interface 14 to the memory system 10 allows data to be written to and read from the system 10 in fixed-size units called sectors, each containing 512 bytes of data, which can be randomly accessed. Each sector is identified by a logical address which in this case is a sequential Logical Block Address (LBA).

In the present arrangement data may be written to a sector even if the sector already contains data. The protocols at the logical interface 14 can, in this case, support, read or write access to the system 10 in multi-sector blocks of logically contiguous sector addresses, these protocols conform to industry standards such as ATA, CompactFlash, or MultiMediaCard thus allowing the memory system 10 to be interchangeable between different host systems and not limited to use with host 12.

The physical interface 18 from controller 16 to Flash Memory 20 allows data to be written to and read from Flash memory 20 in fixed-size units which in this case are called physical sectors and each of which can be



accessed randomly with each typically having sufficient capacity for 512 bytes of data from the host system plus 16 bytes of overhead data appended by the controller 16. Each physical sector is identified by a physical sector address, which normally has separate components which respectively identify the Flash chip within the memory subsystem, the Flash block within the Flash chip, and the physical sector within the Flash block of the memory 20 to which the physical sector is written.

Within the system 10 shown, data may only be written to a physical sector if the sector has previously been erased. The Flash memory 20 is erased in response to a command at the physical interface in units of a Flash block, which typically contains 32 physical sectors. The relative times for performing operations within the Flash system 10 to read a physical sector, program a physical sector, and erase a Flash block are typically in the ratio 1 : 20 : 200.

In the arrangement of Figure 2 the controller 16 is a Cyclic Storage controller which is a Flash media management controller in which a method of ensuring uniformity of distribution of use is implemented wherein the media management algorithms which implement this method are implemented as firmware by a processor within the controller.

With reference to Figure 3 there is shown optimized hardware architecture which is defined for the Cyclic Storage controller 16. In this case the controller hardware is a dedicated architecture in a separate integrated circuit.

The controller 16 comprises host interface control block 22, microprocessor 24, flash interface control block 26, ROM 28, SRAM 30 and expansion port 32, each of these being inter-connected by memory access control bus 34.

Cyclic Storage Flash media management algorithms are implemented by firmware running on microprocessor 24 and the controller 16 is responsible for all Flash media management functions and for the characteristics of the logical interface 14 presented to host 12.

The host interface control block 22 provides the path for data flow to and from host system 12 via logical interface 14.

As, in this case, the controller 16 is in the form of a dedicated integrated circuit the host interface control block 22 provides logical interface 14 which conforms to an industry standard protocol as well as a command register and set of taskfile registers which provide the route for the microprocessor 24 to control the logical characteristics of the interface 14.

The host interface control block 22 also allows for a sector of data to be transferred in either direction across the logical interface 14 between to the host system 12 and the controller's SRAM 30 by a direct memory access (DMA) operation without intervention from the microprocessor 24.

The Flash interface control block 26 provides the path for data flow to and from Flash memory 20, and controls all operations which take place in the Flash memory 20. The operations taking place in Flash memory 20 are defined and initiated by the microprocessor 24, which loads parameter and address information to the flash interface control block 26.

The set of operations which typically take place are the transfer of a physical sector to Flash memory 20, the transfer of a physical sector from Flash memory 20, the programming of a physical sector into flash memory 20, the erasing of a Flash block, and the reading of the status of flash memory 20.

Similarly a physical sector of data may be transferred in either direction across the physical interface 16 between the Flash memory 20 and the controller's SRAM 30 by DMA operations without intervention from the microprocessor 24. The

organization of the 512 bytes of host data and 16 bytes of overhead data within a physical sector which is transferred to Flash memory 20 is determined within the Flash interface control block 26, under the control of parameters loaded by the microprocessor 24.

The Flash interface control block 26 also generates a 12-byte error correcting code (ECC) which is transferred to Flash memory 20 and programmed as overhead data within each physical sector, and which is also verified when a physical sector is transferred from Flash memory 20.

The microprocessor 24 controls the flow of data sectors through the memory access control bus, or datapath, 34 or of the controller 16, implements the Flash media management algorithms which define the sector, controls data storage organisation in the Flash memory 20, and defines the characteristics of the logical interface 14 to host system 12. In this case the microprocessor 24 is a 32-bit RISC processor.

The memory access control bus 34 allows transfer of information between the microprocessor 24, host interface control block 22, and the Flash interface control blocks 16, as well as between the host interface control block 22, the flash interface control block 26 and a memory block 30.

The microprocessor 24, host interface control block 22, and Flash interface control block 26 may each be the master for a transaction on the memory access control bus 34. Bus access is granted to requesting masters on a cycle-by-cycle basis.

The SRAM block 30 stores all temporary information within the controller 16, this storing function includes the buffering of sector data and storage of control data structures and variables, as well as firmware code.

The ROM 28 is included in the controller 16 for storage of code for execution by the microprocessor 24, or of information required by other hardware blocks within the controller.

The inclusion in the controller architecture of an expansion port 32 gives access to external hardware functions, RAM or ROM from the memory system 10.

During the operation of the controller all sector data being transferred between the logical interface 14 to host system 12, and the physical interface 18 to Flash memory 20 is buffered in the SRAM 30. Sufficient capacity in the SRAM 30 is allocated for buffering of two sectors of data to allow concurrent transfers of successive sectors at the host and Flash interfaces. Data transfer between the logical host interface 14 and SRAM

30 is performed by DMA with the host interface control block 22 acting as bus master.

Data transfer between the physical Flash interface 18 and SRAM 30 is performed by DMA with the Flash interface control block 26 acting as bus master. Data to be written to sectors in Flash memory 20 is stored in the SRAM memory 30 and is transferred by direct memory access under the control of the Flash interface control block 26 via the physical interface to Flash memory 18. 512 bytes of user data to be written in a sector had previously been supplied by host system 12 via the logical interface 14 and had been transferred by direct memory access under the control of the host interface control block 22 to the SRAM memory 30. Programming of data in a sector in Flash memory 20 is accomplished by the controller 16 by sending an address and command sequence at the physical interface 18, followed by 528 bytes of data plus ECC, followed by a program command code.

The transfer of data for a sector between a host system and the controller's SRAM 30, and between the SRAM 30 and Flash memory, is controlled by firmware running on the microprocessor 24 with the controller 16 being responsible for all Flash media management functions and for the characteristics of the logical interface 14 present to host 12.

As the controller 16 is in the form of a dedicated integrated circuit, the host interface control block 22 provides a logical interface which conforms to an industry standard protocol, and a command register and set of taskfile registers provide the route for the microprocessor 24 to control the logical characteristics of the interface 14. Command, address and parameter information is written to these task file registers by the host 12, and read by the microprocessor 24 for execution of the command. Information is also be written to the registers by the microprocessor 24 for return to the host 12.

In Figure 4 there is illustrated the layered structure of the firmware which performs the Cyclic Storage Flash media management operations. The firmware has three layers, the first being the host interface layer 40, the second layer 42 comprising the sector transfer sequencer 42a and the media management layer 42b and the third being the flash control layer 44.

These three firmware layers 40, 42 and 44 control the transfer of data sectors between the logical interface 14 to host 12 and the physical interface 18 to Flash memory 20. However, the firmware layers do not directly pass data, instead data sectors are transferred by the

hardware blocks of the controller 16 and therefore do not pass through the microprocessor 24.

The host interface layer 40 supports the full command set for the host protocol. It interprets commands at the host interface 14, controls the logical behavior of the interface 14 according to host protocols, executes host commands not associated with the transfer of data, and passes host commands which relate to data in Flash memory to be invoked in the layers below. Examples of such commands are.

Read logical sector (single or multiple),

Write logical sector (single or multiple),

Erase logical sector (single or multiple), as well as other disk formatting and identification commands.

The sector transfer sequencer 42a receives interpreted commands relating to logical data sectors from the host interface layer 40 and thus invokes the Flash media management layer 42b for logical to physical transformation operations, and invokes the Flash control layer 44 for physical sector transfers to or from Flash memory. The sector transfer sequencer 42a also performs sector buffer memory management. Another function of the sequencer 42a is to create a sequence of sector transfers, at the host interface 14 and Flash memory



interface 18, and a sequence of operations in the media management layer 42b, in accordance with the command received from the host 12 and the level of concurrent operations which is configured for the Flash memory 20.

The media management layer 42b performs the logical to physical transformation operations which are required to support the write, read or erasure of a single logical sector. This layer is responsible for the implementation of Cyclic Storage media management algorithms.

The Flash control layer 44 configures the Flash interface control block 26 hardware to execute operations according to calls from the sector transfer sequencer 42a or media management layer 42b.

The media management functions which are implemented within the media management layer 42b of the controller firmware create the logical characteristics of a disk storage device in the memory system 10 which uses Flash semiconductor memory 20 as the physical data storage medium.

The effectiveness of the media management performed by the media management functions of the media management layer 42b is measured by its speed for performing sustained writing of data to the memory system 10, its efficiency in maintaining its level of performance when operating with different file systems, and in this case,

in host 12, and the long-term reliability of the Flash memory 20.

Data write speed is defined as the speed which can be sustained when writing a large volume of contiguous data to the memory system 10. In some cases, when the sustained data write rate of a memory system is being tested, the volume of data to be written may exceed the capacity of the memory system 10 and therefore logical addresses may be repeated.

Sustained write speed is determined by the sector data transfer speed at the logical interface 14 to the host 12, and the physical interface 18 to Flash memory 20, as well as the overhead percentage of accesses to Flash memory 20 at the physical interface 18 for Flash page read and write operations and Flash block erase operations which are not directly associated with storage of data sectors written by the host 12 at the logical interface 14. In this case the control data structures and algorithms which are employed should ensure that access to Flash memory 20 for control functions is required at a much lower frequency than for host sector write. The sustained write speed is also determined by the processing time within the controller 16 for media management operations, and the page read and program times, and block erase times within the Flash memory 20.

In order for the memory system to operate efficiently when having file systems with different characteristics, the Media management algorithms for the organization of host data and control data structures on Flash memory 20 are appropriately defined and data write performance is maintained in each environment.

In a first embodiment, the file systems implementing the MS-DOS standard are provided with at least one of the following characteristics: the host 12 writing data sectors in clusters using multiple sector write commands; the host 12 writing data sectors using single sector write commands; the host 12 writing some sectors with single sector write commands in an address space which is shared with clustered file data; the host 12 writing non-contiguous sectors for MS-DOS directory and FAT entries with single sector write commands; the host 12 writing non-contiguous sectors for MS-DOS directory and FAT entries interspersed with contiguous sectors for file data; and/or

the host may rewrite sectors for MS-DOS directory and FAT entries on a frequent basis.

It is a feature of flash memory, and in this case the Flash memory 20 of the memory system 10, that it has a wear-out mechanism within the physical structure of its cells whereby a block of flash memory may experience

failure after a cumulative number of operations. Typically, this is in the range of 100,000 to 1,000,000 program/erase cycles. In light of this the cyclic storage controller 16 of the present arrangement implements a process of wear-leveling to ensure that "hot-spots" do not occur in the physical address space of the Flash memory 20 and that utilization of Flash blocks is uniformly distributed over a prolonged period of operation.

Cyclic Storage media management algorithms are implemented within memory system 10 and perform the Media management operation of the physical Flash memory 20 within the system 10. The cyclic storage media management algorithms comprise four separate algorithms, namely the Data Write algorithm which controls the location for writing host information to, the Block Erase algorithm which controls erasure of areas of Flash memory 20 containing obsolete information, the Block Sequencing algorithm which controls the sequence of use of Flash blocks for storing information, and the Address Translation algorithm which controls the mapping of host logical addresses to physical memory addresses.

The method of Cyclic Storage media management implemented by these algorithms embodies the principle that data is written at physical sector locations in

Flash memory 20 which follow the same order as the sequence in which the data is written. This is achieved by writing each logical data sector at a physical sector position defined by a cyclic write pointer.

A schematic representation of the write operation of the cyclic storage media management method is shown in Figure 5A. The write pointer, in this case data write pointer (DWP) 46 moves sequentially through the sector positions of Flash block X in Flash memory 20, and continues through the chain of blocks Y and Z in a manner defined by the block sequencing algorithm. Each block X, Y and Z is a physical structure in Flash memory 20 which, in this case, comprises 32 sector locations which can be erased in a single operation.

As is illustrated in Figure 5B logical data sectors are generally written in files by a file system in the host 12, and the Cyclic Storage Data Write Algorithm locates the first sector of a file at the next available physical sector position following the last sector of the preceding file. When a file is written by host 12 using logical sectors for which valid data already exists in the device, the previous versions of the sectors become obsolete and the blocks containing them are erased according to the Block Erase Algorithm. In order to erase a block containing obsolete file sectors it is, in

some cases necessary to relocate some valid sectors of another file. This generally occurs when a block contains sectors of the head of a file, as well as sectors with unrelated logical addresses from the tail of a different file.

A second write pointer in this case data relocate pointer DRP 47 is used for writing relocated sectors in order to avoid sectors of one file fragmenting a block containing sectors of another file. The use of a separate relocation pointer significantly reduces the fragmentation of blocks containing a file, leading to minimum requirement for sector relocation and consequent maximum file write performance.

A host file system is used which also writes sectors containing system information, such as directory or FAT sectors in the DOS file system, and these are generally written immediately before and after a group of sectors forming a file. A separate system pointer, system write pointer SWP 48 is used for this host file system in order to define the physical write location for system sectors, which are identified by their logical address, in order to separate system sectors from file data sectors and avoid them being treated in the same way. This avoids a small group of system sectors being "sandwiched" between the tail of one file and the head of another. These

system sectors contain information about many files, and are generally re-written much more frequently than data for a file. "Sandwiched" system sectors would cause frequent relocation of file data sectors and thus the use of system pointer SWP 48 minimises the requirement for data sector relocation and maximises file write performance.

A fourth pointer, system relocate pointer SRP 49 is used for relocation of system sectors, analogous to the relocation pointer DRP 47 for file data sectors.

To summarise, the four write pointers are:

Data write pointer, DWP 46 which is used to define the physical location for writing file data sectors transmitted by a host system;

System write pointer, SWP 48 which is used to define the physical location for writing system sectors transmitted by a host system wherein system sectors are identified by their logical address, in accordance with the characteristics of the host file system in use;

Data relocation pointer, DRP 47 which is used to define the physical location for writing file data sectors which must occasionally be relocated prior to a block erasure for recovery of capacity occupied by obsolete file data sectors; and

System relocation pointer, SRP 49 which is used to define the physical location for writing system sectors which are being relocated prior to a block erasure for recovery of capacity occupied by obsolete system sectors.

A block must contain data associated with only a single write pointer and this results in four separate chains of blocks existing, one for each write pointer. However, the same write and relocation algorithms of the cyclic storage algorithms apply to each write pointer 46, 47, 48 and 49.

This scheme for locating a sector to be written at the first available location following the preceding sector, combined with usage of multiple write pointers, is fully flexible, and provides high performance and total compatibility for all host write configurations, including single sector data and data in clusters of any size.

However, the Cyclic Storage media management method is defined not to allow the existence of a large number of obsolete data sectors and nor to implement background operations for functions such as garbage collection. Typically only two blocks containing obsolete sectors is allowed to exist for each of the Data Write Pointer DWP 46 and System Write Pointer SWP 48, and block erasure is



performed as a foreground operation during sector write sequences.

This method of management means that the logical capacity of the flash memory does not have to be reduced to allow for the existence of a large volume of obsolete data, the data integrity is significantly improved by the absence of background operations, which are susceptible to interruption by power-down initiated by the host; and the pauses in data write sequences are short because erase operations are required for only a single block at a time.

If an obsolete data sector is created in a new block associated with either of the write pointers, then the existing "obsolete block" is eliminated by erasure, following sector relocation within the blocks if required.

Erase sector commands sent from a host 12 are supported by marking the target sector as obsolete, and allowing its erasure to follow according to the Block Erasure algorithm.

The Cyclic Storage block sequencing algorithm determines the sequence in which blocks within the flash memory 20 are used for the writing of new or relocated data, and is therefore responsible for ensuring that no

block experiences a number of write/erase cycles which exceeds the endurance limit specified for the Flash memory system 10 which is being used.

When a logical sector is written by the host, any previous version which exists in the memorys system is treated as obsolete data. The block erase algorithm ensures that blocks which contain obsolete data sectors are erased immediately, to allow recovery of the capacity occupied by these sectors. The physical memory capacity of the system 10 is therefore occupied by valid data for logical sectors written by the host, plus a small number of proprietary Cyclic Storage control data structures and a number of erased blocks. Immediately after initial formatting, of the flash memory 20 the capacity of the memory 20 consists almost entirely of erased blocks. When the host 12 has written at least once to all sectors in its logical address space, the device is considered to be logically full and its physical capacity is occupied almost entirely by valid data sectors, with a small number of erased blocks maintained for correct device operation. An increased number of erased blocks will be created only if the host 12 executes commands to erase logical sectors.

Erased blocks which are allocated for use by one of the write pointers, or for storage of control data

structures are taken from a pool of available erased blocks. A block is never erased in response to a need to perform a write operation to that specific block, the block sequencing algorithm determines the order of allocation for data write operations of blocks in the erased pool. The next available block according to the algorithm is allocated, independent of whether the requirement is for use by one of the write pointers or for a control data structure.

The implementation of these algorithms which perform the cyclic storage media management allows increased system flexibility by operating on individual sectors of the flash memory 20 and separately tracking the logical to physical address mapping of every sector in its logical address space. A sector address table is maintained in the Flash memory 20 which contains the physical address for every logical sector. In addition, every sector is written with a header containing its logical address, providing a means of verifying sector identity and ensuring maximum data integrity.

The data write algorithm, with its use of cyclic write pointers, provides the capability for tracking the sequence of sector writing using the logical addresses in the headers of sectors in sequential physical positions. This feature provides total data security even when the

logical to physical address mapping records for recently written sectors are temporarily held in volatile controller memory SRAM 30 and not in Flash memory 20. Such temporary records can be reconstructed from the data sectors in Flash memory 20 when a system 10 in which the Cyclic Storage algorithms are implemented is initialised. It is therefore possible for the sector address table in Flash memory 20 to be updated on an infrequent basis, leading to a low percentage of overhead write operations for control data and a high sustained data write rate.

In Figure 6 there is shown a schematic representation of the address translation process which uses a three level hierarchy of mapping structures 50 which is performed in the memory system 10.

The three levels of the hierarchy are the sector address table 52, the temporary sector address table 54 and the sector address record 56.

The top level of the hierarchy of the mapping structures is the sector address table 52, which is a master table containing a physical address for every logical sector stored in the system 10 and which is stored in Flash memory 20. Structures in the two lower levels of the hierarchy 54 and 56 provide the means for reducing the frequency at which write operations must occur to the sector address table.

The sector address record 56 is a list stored in the controller's volatile memory SRAM 30 of logically contiguous sectors which have been written to system 10. This list allows the physical address of any logical sector which it contains to be determined without need for access to Flash memory 20. It may also be reconstructed during device initialisation from the sequence of recently-written sectors which may be traced in the Flash memory 20. The intermediate temporary sector address table 54 is contained in Flash memory 20 and is updated with the contents of the sector address record 56 when the list becomes full. The intermediate temporary sector address table 54 is in the same format as the sector address table 52, and allows physical address data updates to specific blocks of the sector address table 52 to be accumulated to allow a more efficient table write process to be performed. The temporary table 54 allows the physical address of any logical sector contained in it to be determined without need for access to the sector address table 52.

This hierarchy of mapping structures 50 is maintained with an infrequent requirement for write operations to Flash memory and efficiently supports logical to physical address translation in such a way that total security of sector address information is

provided, even if electrical power is unpredictably removed from the system 10.

The data structures required to support the Cyclic Storage media management algorithms are stored principally in Flash memory 20 together with the host data sectors, with only a very limited amount of control data being held temporarily in the control processor's volatile RAM 30. Information held in the volatile memory 30 is non-critical, and can be reconstructed from Flash memory 20 if the power supply is interrupted.

The controller 16 in Flash memory system 10 as described above, may operate on only one array within the Flash memory 20 at a time. Each array is a group of Flash memory storage cells within which only a single sector program operation or block erase operation may be performed at any one time. In this case the array is a complete Flash chip. The controller is designed to be capable of performing program operations concurrently on sectors within different arrays or erase operations concurrently on blocks within different arrays. The controller 16 can address, program and check current status of any array within the Flash memory 20 independently from others.

Each sector is a unit of physical storage in Flash memory 20 which is programmed in a single operation. In

the present arrangement, which comprises NAND Flash memory chips, a sector equivalent to a page within the Flash array and has a capacity of 528 bytes. In this case, the each Flash chip is considered to comprise four arrays, each of which can be programmed with one sector at any time.

The scheduling of transfer ie the ordering of sector data is controlled by the sector transfer sequencer block 42a shown in Figure 4 and is explained in greater detail with reference to Figures 7 to 11. The transfer of data at the host interface 14 is independent of transfer of data at the physical interface to Flash memory 18 of memory system 10, and the burst transfer rate at the host interface is determined by the host 12. Several different methods of scheduling the transfer of sector data may be implemented by sector transfer sequencer firmware, depending on the way in which blocks and pages in the Flash memory 20 are addressed by the controller 16. The methods described assume that sector data is supplied by the host and stored in SRAM 30 at a rate which is sufficient to supply sector data for transfer to Flash memory 20 as described.

With reference to Figure 7, there is shown a first embodiment of a method of scheduling the transfer of sector data which the controller may use to address

blocks and pages within memory system 10 wherein Flash memory 20 comprises four Flash arrays 0, 1, 2 and 3, where the controller 16 is required to initiate concurrent page program or block erase operations in two arrays simultaneously. The arrays are linked in pairs, and the corresponding blocks 0 with the same addresses within the linked arrays are treated as a single virtual block 0. As shown, block 0 in Flash array 0 is linked with block 0 in Flash array 1 to form virtual block 0. The N blocks in each of Flash arrays 0 and 1 are linked to form N virtual blocks, labelled 0 to N-1, and the N blocks in each of Flash arrays 3 and 2 are linked to form a further N virtual blocks, labelled N to 2N-1. The order of writing sectors within each virtual block is determined by the movement of the write pointer, which alternates between the constituent blocks as it moves sequentially through the sectors in the virtual block.

In Figure 8 there is shown a second embodiment of a method which the controller may use to address blocks and pages within memory system 10 wherein Flash memory 20 comprises four Flash arrays, where it is necessary for the controller 16 to initiate concurrent page program or block erase operations in four arrays. All four arrays are linked, and the corresponding blocks with the same addresses within each of the linked arrays are treated as



a single virtual block. As can be seen, blocks 0 in Flash arrays 0 to 3 are linked to form virtual block 0. The N blocks in each of Flash arrays 0 to 3 are linked to form N virtual blocks, labelled 0 to N-1. The order of writing sectors within a virtual block is determined by the movement of the write pointer, which moves through the corresponding sectors in blocks 0 to 3 and then increments to the next sector in block 0, as it moves sequentially through the sectors in the virtual block.

The blocks within the individual Flash arrays which are linked to form a virtual block may themselves comprise multiple smaller adjacent physical blocks which are stacked together.

Program operations may be performed substantially concurrently on one sector from each of the constituent blocks forming a virtual block.

With reference to Figure 9, there is shown a third embodiment of a method of concurrently programming sectors 0, 1, 2, and 3 which are shown in Flash arrays 0, 1, 2 and 3 of Figure 8. Data for sector 0 is transferred byte serially to Flash array 0 across the physical interface to Flash memory 20, and then a program command is sent by the controller 12 to Flash array 0 to initiate the program operation. Whilst sector 0 is being programmed, the controller 12 transfers data for sector 1

to Flash array 1 and initiates a program operation for it. The same is done for sectors 2 and 3. Sectors 0 to 3 are programmed in Flash arrays 0 to 3 substantially concurrently with each other, and the speed of transferring and programming data to sectors in the Flash memory is much higher than can be achieved by programming only one Flash array at a time. When the program operations in Flash arrays 0 to 3 have all completed, the process is repeated for sectors 4 to 7. A shared busy/ready line from the Flash arrays can be used to signal when all arrays have completed programming sectors 0 to 3 and when there is no Flash array active. However, the status of all the arrays can alternatively be polled independently.

In Figure 10 there is shown a fourth embodiment of a sequence for transferring sector data to and initiating programming operations in Flash arrays 0 to 3. The sequence described for sectors 0 to 3 with reference to Figure 9 is performed, but upon the completion of the programming operation in a Flash array, sector data is immediately transferred for the following programming operation in that array. The status of each array is polled independently to find when an operation in the array has completed. Alternatively, independent ready/busy signals from every array can be used. This

increased pipelining of the sector data transfer, and sector programming provides further increased speed for writing sector data in the Flash memory.

Each of these methods detailed in the above described embodiment may be used for writing sector data which is being relocated from another sector in Flash memory, as well as sector data which has been supplied by a host system.

The order of sectors being concurrently programmed in different Flash arrays need not follow the order shown in Figure 8, that is, sequential order need not be used. It is possible to program any sector from a Flash array concurrently with any other sector from another array, provided that no two sectors from the same array are used. For example, it would be possible to transfer and then program a group of four sectors in Figure 8 in the order sector 10, sector 3, sector 1, then sector 4. However, the use of a cyclic write pointer which moves sequentially through the addresses of a virtual block means that it is most common for sector addresses to be in sequential order. The first sector of a group of four for which data is being concurrently transferred and programmed need not be located in Flash array 0. The sectors may, for example, be transferred in the order sector 2, sector 3, sector 4, and sector 5.

The write time for a cluster of sectors can be expressed as a function of the transfer time to Flash memory 20 for sector data and programming time for a sector in Flash memory 20. The programming time is typically 200 microseconds, and is much longer than transfer time, which is typically about 30 microseconds. The time associated with flash chip addressing and initiation of data transfer and programming by the controller is usually not significant. For the example shown in Figure 9, cluster write time is given by

$$\begin{aligned} \text{Cluster Write Time} &= 8 * \text{Sector data transfer time} \\ &+ 2 * \text{Programming time.} \end{aligned}$$

For the example shown in Figure 10, cluster write time is given by

$$\begin{aligned} \text{Cluster Write Time} &= 5 * \text{Sector data transfer time} \\ &+ 2 * \text{Programming time.} \end{aligned}$$

As detailed above, all 4 flash memory arrays are being accessed, however this results in the electrical current level being high (= 4\*array current) as well as performance is on the maximum.

With reference to Figure 11, there is illustrated a first embodiment of a method of allowing the limiting of the number of flash memory arrays accessed at a time in order to control the current. As can be seen, Figure 11 illustrates a pipelined sequence for transferring sector

data to and initiating programming operations in flash arrays 0 to 3 but the maximum number of active flash memory arrays is limited to three. The sequence described for sectors 0 to 2 with reference to Figures 9 and 10 is performed, but then the number of active arrays is three and therefore the controller 16 waits for at least one of the arrays to be completed, in this case chip 0. Whenever the number of active arrays is lower than three, sector data is immediately transferred for the following programming operation in the next chip. In other words, before accessing any flash array the number of currently active arrays is always checked against the allowed limit which in this case is three. Thus, the electrical current level is regulated.

For the example illustrated by Figure 11 the cluster time will be

$$\text{Cluster Write Time} = 4 * \text{Sector data transfer time} + 3 * \text{Programming time}.$$

Similarly, the number of active arrays can be limited to two or one. For the same case of a 4-way interleaved memory system where the active array limit is two (not shown) the cluster time will be

$$\text{Cluster Write Time} = 5 * \text{Sector data transfer time} + 4 * \text{Programming time}.$$

A more complex method of performing the control of current can be used when the flash memory 20 of the system 10 has different electrical parameters for different flash operations, ie the electrical parameters for read, transfer, programming and erase operations are all different. For simplicity the arrays are as before, programmed in numerical order 0 to 3. The status of every array is polled independently to find a time when the chip completed. Otherwise, independent ready/busy signals from every array can be used. This method uses the same method of pipelining but creates extra on-purpose delay in order to limit the electrical level.

The flash access control combined with the pipelining provides a very efficient usage of flash memory performance and still gives a high speed for writing sector data in the flash memory while limiting the electrical power. The method allows a defined number of active arrays, in this case three, most of the time. The same method can be applied to any other flash operation like read and erase.

The maximum number of active memory arrays can be flexibly changed/programmed by the host to define the ratio between write performance and electrical current level. Some hosts could implement this by enabling standard-defined power management features and defining a

level of compromise between power consumption and performance. However, some host systems may prefer slow memory devices with low electrical current levels, and in some cases even the same host system may prefer different combinations of performance and power consumption in different operating modes. Each of these arrangements can be catered for using the method described above.

Various modifications may be made to the arrangements as hereinbefore described without departing from the scope of the invention. For example, a system which incorporates a flash disk device may be physically partitioned in several ways, according to the system architecture, however, all systems generally conform to the structure described herein before. For example the flash memory 20 is shown in Figure 2 as being part of a memory system 10, however, it may alternatively be on a removable card and may connect to a host system via a logical interface 14 which as before conforms to industry standard protocols. Examples of such industry standards being PCMCIA ATA, CompactFlash and MultiMediaCard. In such an arrangement the controller may be on a removable card in which case the controller is typically a single integrated circuit. The Flash memory 10 may consist of one or more integrated circuits and the controller may

be integrated on the same integrated circuit as the Flash memory.

It could also be the case that the host and the flash system may be physically partitioned such that only the Flash memory is on a removable card, which has a physical interface to the host system. A hierarchy of this arrangement is shown in Figure 12. An example of such a removable Flash memory card is SmartMedia. The controller is located within the host system 11 and may take the form of an integrated circuit, or of firmware which is executed by a processor within the host system.

Alternatively the method of the present invention may be implemented in an embedded memory system which is not physically removable from a host system. Such a system may have the same partitioning as is used for a memory system on a removable card, with the controller being in the form of an integrated circuit and with a logical interface conforming to industry standard protocols. However, the controller may also be integrated with other functions within the host system.

In the arrangement described, each sector is identified by a LBA, however, it may also be identified by an address in the Cylinder/Head/Sector (CHS) format originally used with magnetic disk devices. Also in the described arrangement the controller hardware is



dedicated architecture in a separate integrated circuit, however, elements of the controller hardware, such as the microprocessor, may be shared with other functions within the host system. Additionally the cyclic storage management algorithm may be implemented in a microprocessor within the host system or the process may be performed via standard microprocessor input/output ports without any dedicated controller hardware. If the controller is part of an embedded memory system and shares its microprocessor with other functions of a host system, the logical interface for the control of the memory system may be implemented directly within firmware executed by the processor, this means that hardware registers may be eliminated and variables may be passed directly to a controller function which maybe called a host function withing the firmware code.

In the flash memory system described previously, data transfer between the host or flash interfaces and the SRAM are performed by DMA however in an alternative embodiment a separate memory block could be used exclusively for buffering sector data. Typically this memory block could be a dual port RAM, with ports allocated independent access by the host interface control block and the flash interface control block.

In the described arrangement the memory blocks into which the memory sectors were arranged were described as being a physical structure within the flash memory comprising 16 sector locations, however it is also possible that these memory blocks comprise 32 flash locations. Also the memory blocks can alternatively be virtual blocks comprising physical blocks distributed across multiple flash chips or multiple independent arrays within the same chip which are erased in a single operation by the controller. Where a virtual block comprises  $M$  physical blocks, each with capacity for  $N$  sectors, the virtual block has capacity for  $M*N$  sectors. A virtual block is treated in exactly the same way as a physical block by the cyclic storage media management algorithms.

It should also be noted that the ROM and expansion port of the controller of the memory system are optional features and need not be included.

Furthermore, each array in the flash memory is described previously as being a complete flash chip, however, it is also the case that each array may be a constituent part of a chip, as some Flash chips such as some 512Mbit NAND flash designs incorporate multiple arrays within a chip and separate sector program operations may be independently started in different

arrays within the chip. Also in the description, pages within the flash array have been described as being equivalent to a sector, however in some AND flash memory chips a page may comprise four sectors and have a capacity of 2112 bytes, in each case the page is programmed in a single operation. Additionally each group of sector data has been described as being the first four sector data of a file, however it may alternatively be a file fragment. Also the host system can write data to the memory system in units of a cluster wherein each cluster will be treated as the controller as an integral number of groups, as opposed to the data being written to the memory system as single sectors.

The present invention can be implemented by a computer program operating on a microprocessor of a memory system. An aspect of the present invention thus provides a storage medium storing processor implementable instructions for controlling a processor to carry out the method as hereinabove described.

CLAIMS

1. A method for use in a memory system having a non-volatile memory and a controller for limiting the number of non-volatile memory arrays, within the non-volatile memory, from a plurality of available arrays accessed at one time, the method comprising:  
  
implementing a pipelining sequence for transferring data to and from the non-volatile memory arrays; and  
  
limiting the number of active arrays operating at one time, the arrangement being such that the controller waits for the at least one of the arrays to complete before initiating the transfer to and from a further array,  
  
thereby controlling flash memory accesses in a memory system configured to use concurrent operation in different flash memory arrays and allowing performance to be changed easily in systems supporting concurrent flash operations in different Flash arrays.
2. A non-volatile memory system comprising:  
  
a non-volatile memory incorporating a plurality of memory arrays; and  
  
a controller including,  
  
means for implementing a pipelining sequence for transferring data to and from the non-volatile memory arrays; and  
  
means for limiting the number of active arrays operating at one time, the arrangement being such that the controller waits for the at least one of the arrays to complete before initiating the transfer to and from a further array,  
  
thereby controlling flash memory accesses in a memory system configured to use concurrent operation in different flash memory arrays and allowing performance to be changed easily in systems supporting concurrent flash operations in different Flash arrays.

3. A storage medium storing processor implementable instructions for controlling a processor to carry out the method of claim 1.

5 4. A method substantially as hereinbefore described with reference to and as shown in any of Figures 2 to 12 of the accompanying drawings.

0 5. A system substantially as hereinbefore described with reference to and as shown in any of the accompanying drawings.

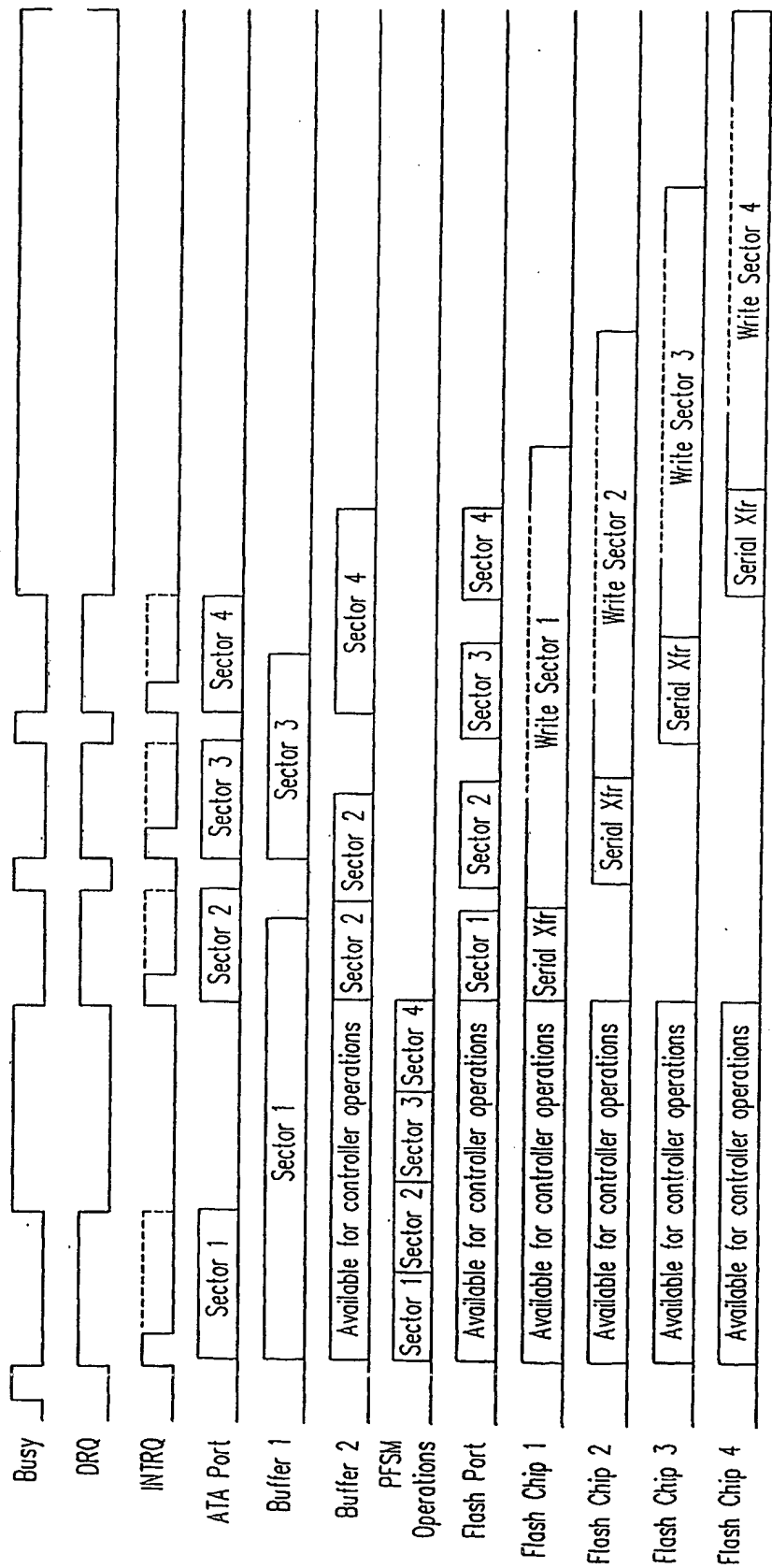


FIG. 1 (Prior Art)

2/9

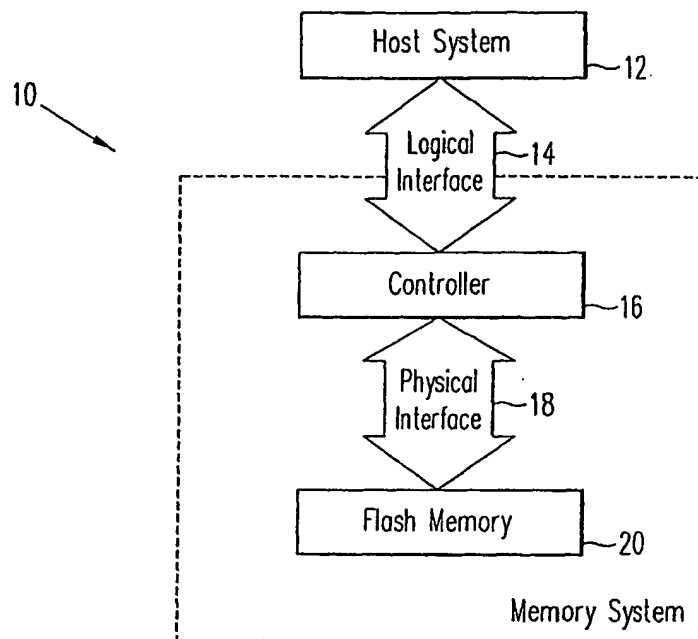


FIG. 2

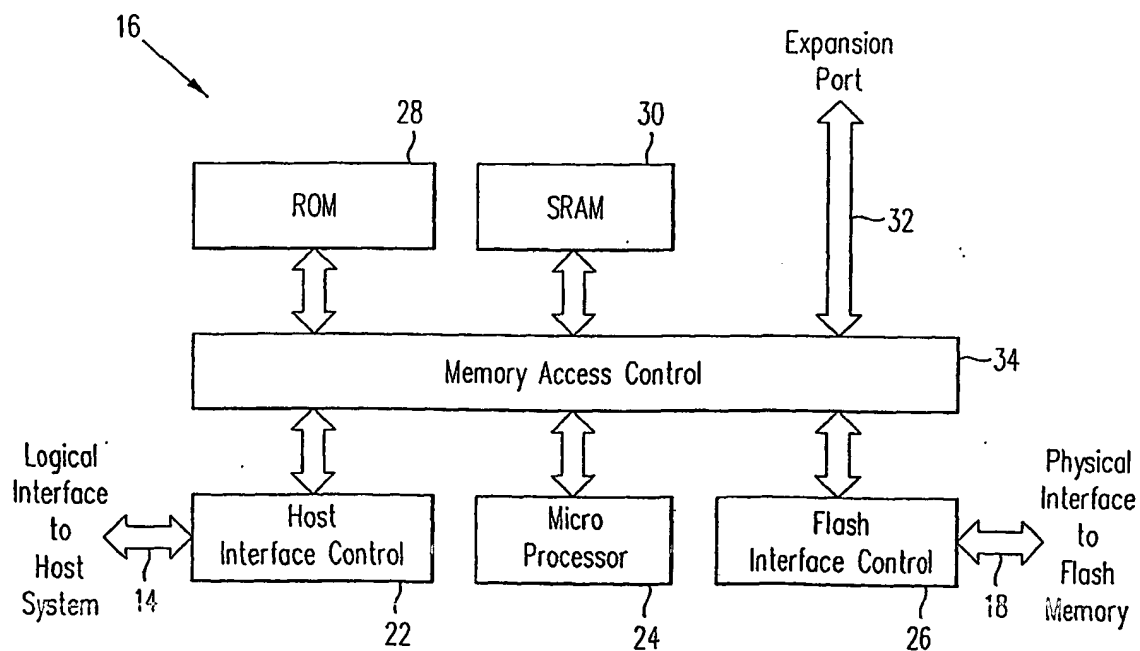


FIG. 3

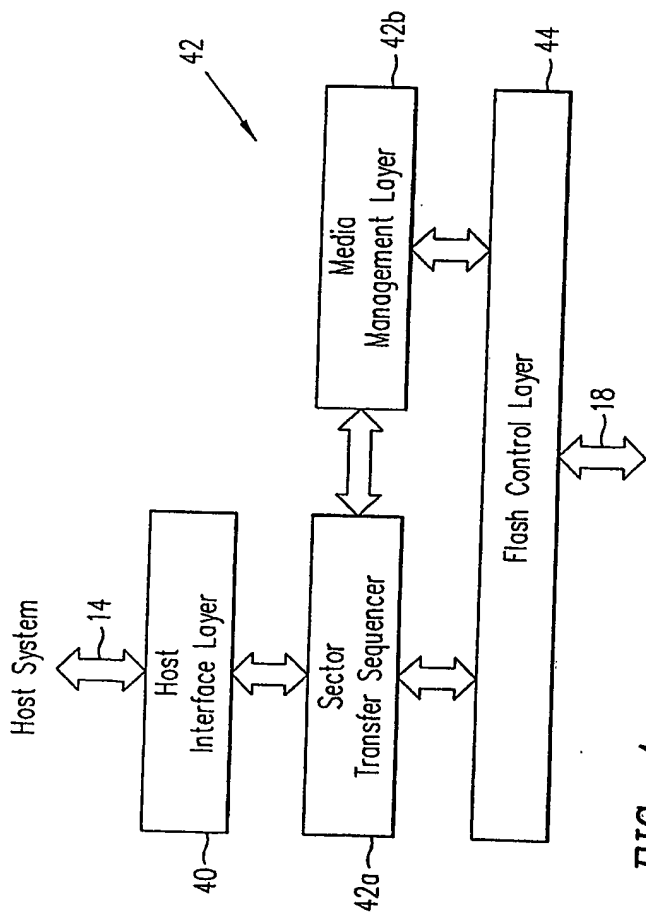


FIG. 4

Flash Interface Control Hardware

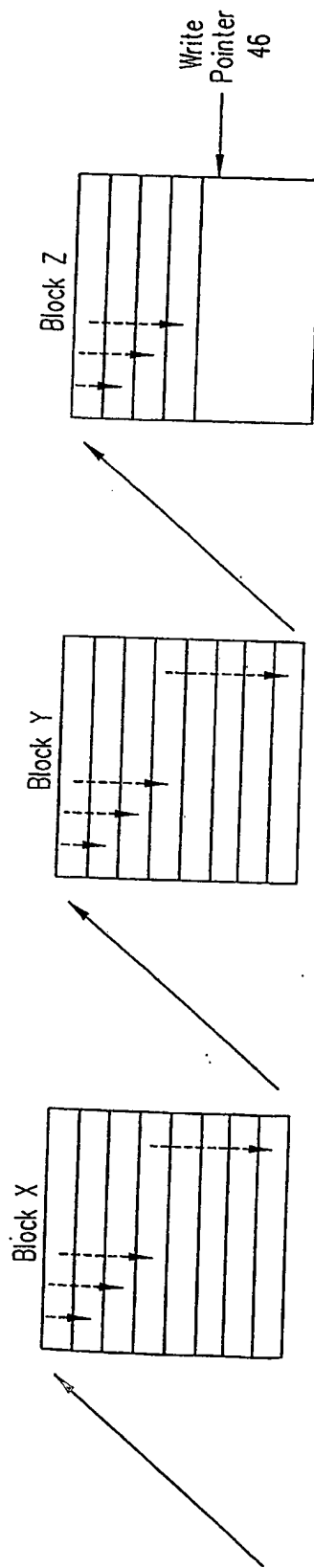


FIG. 5A



4/9

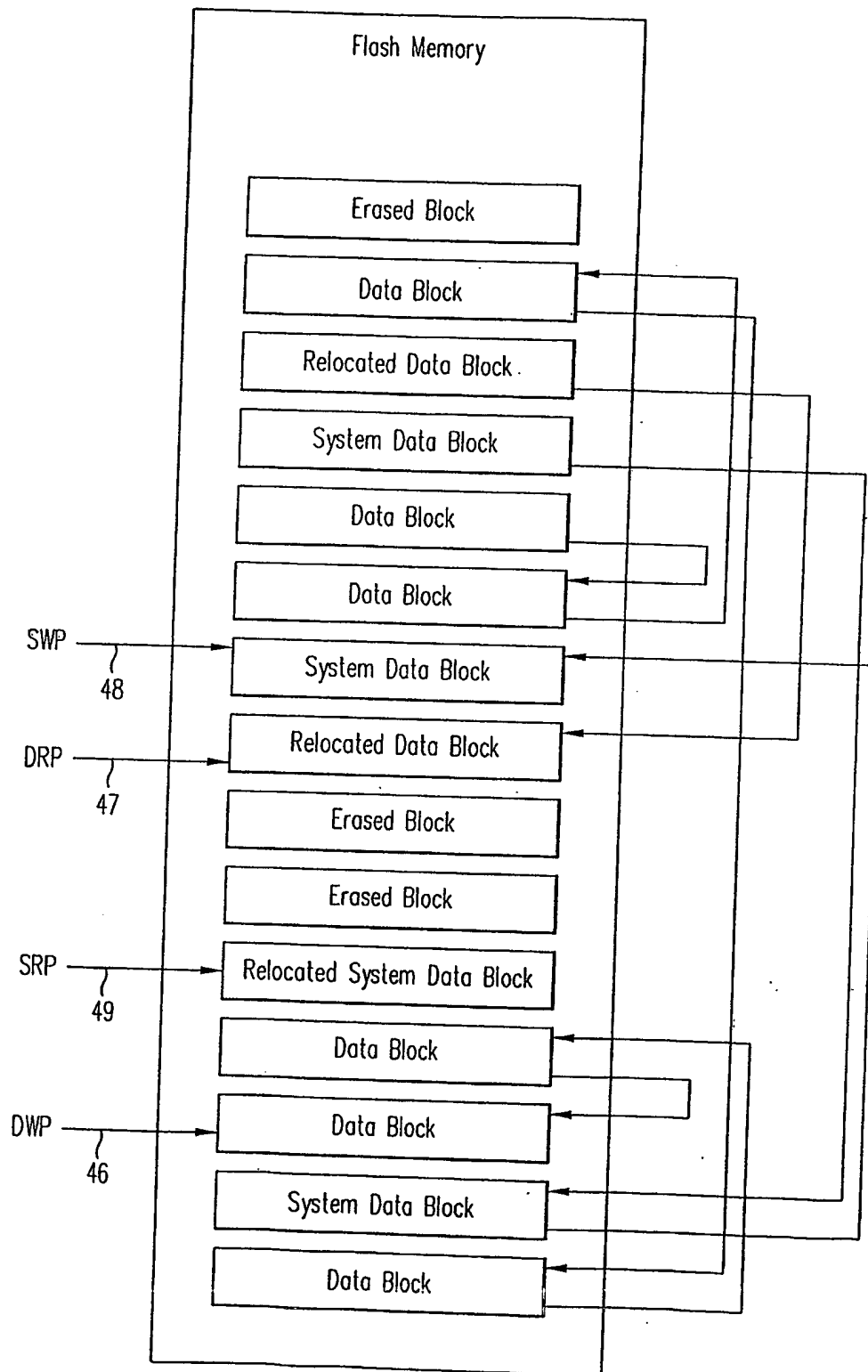


FIG. 5B

5/9

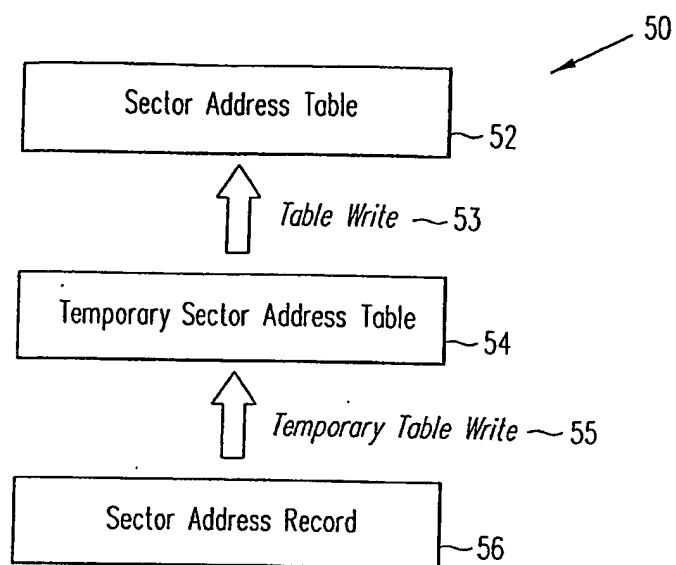


FIG. 6

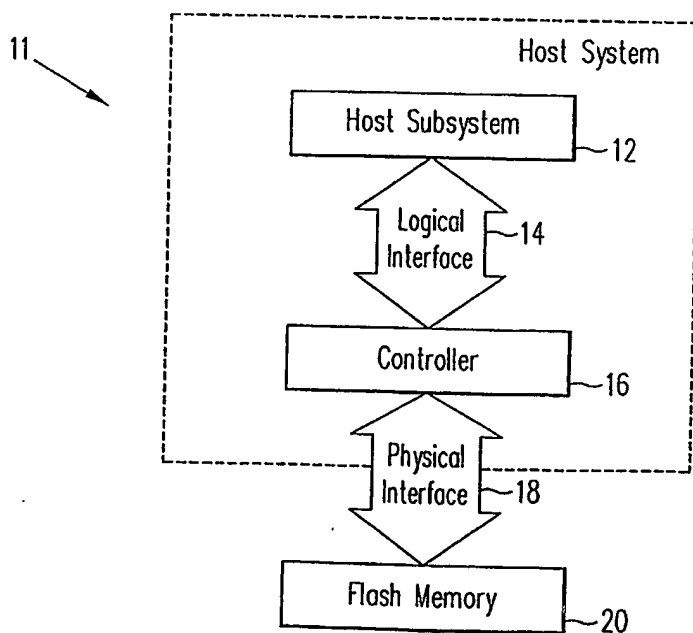


FIG. 12

6/9

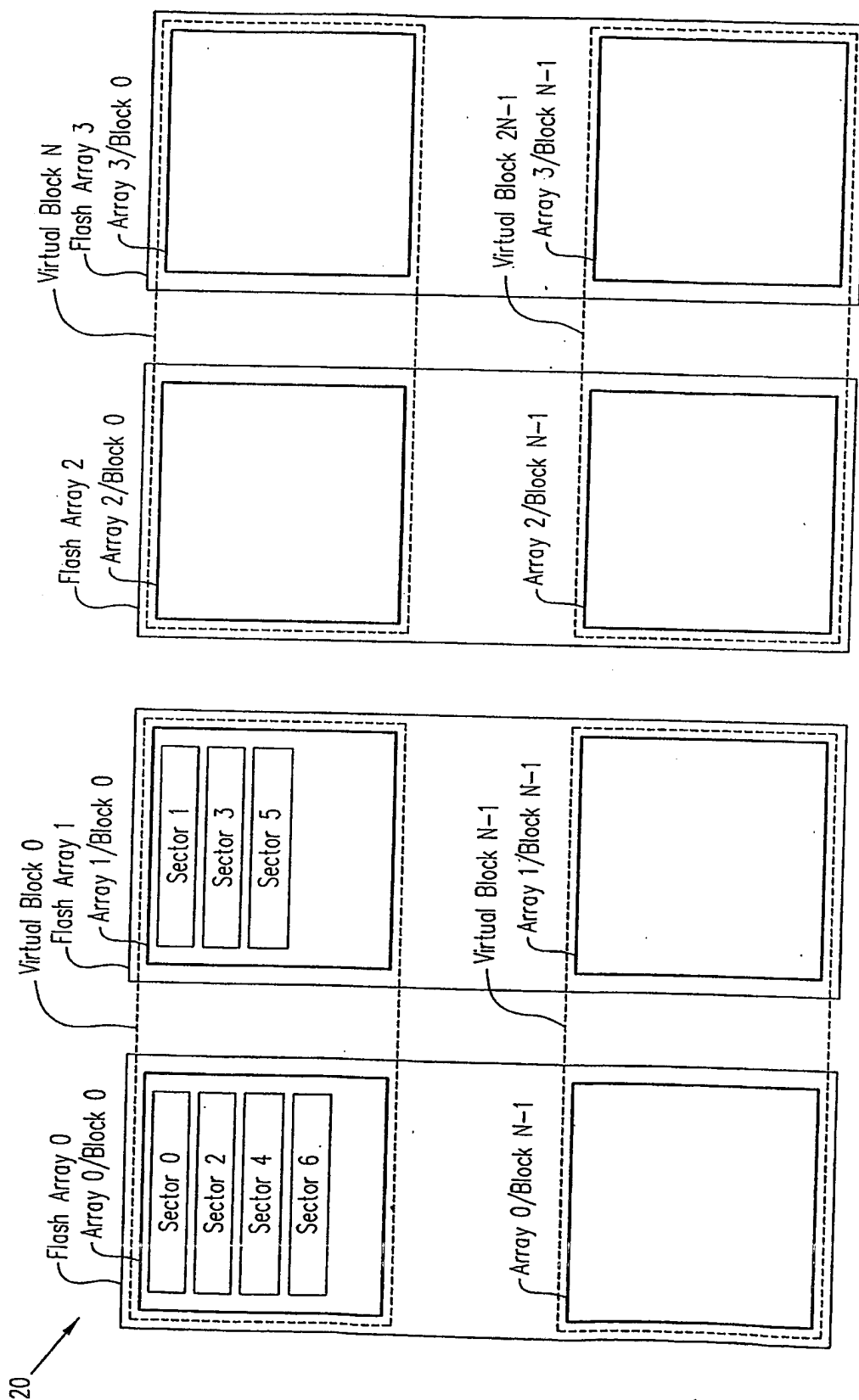


FIG. 7

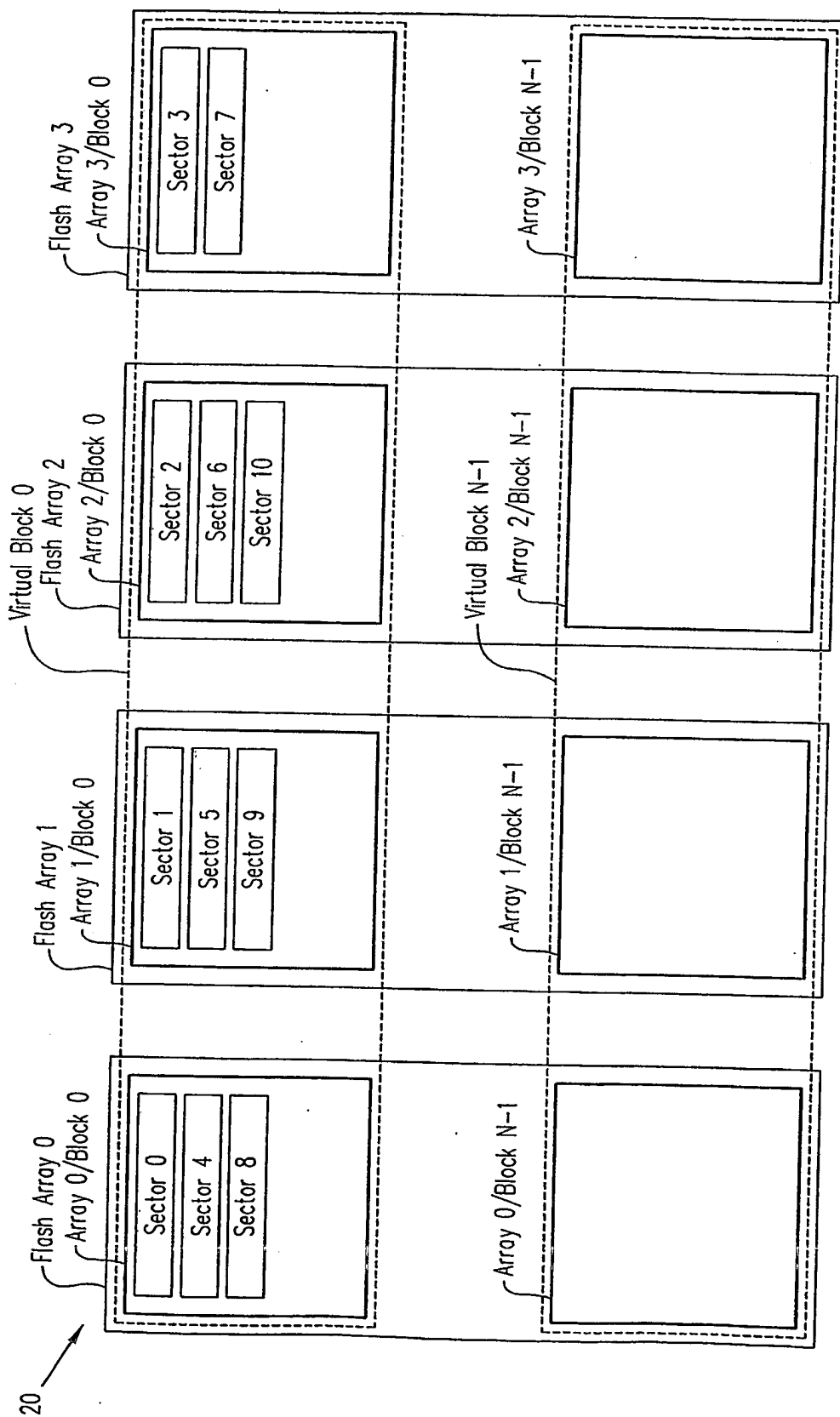


FIG. 8

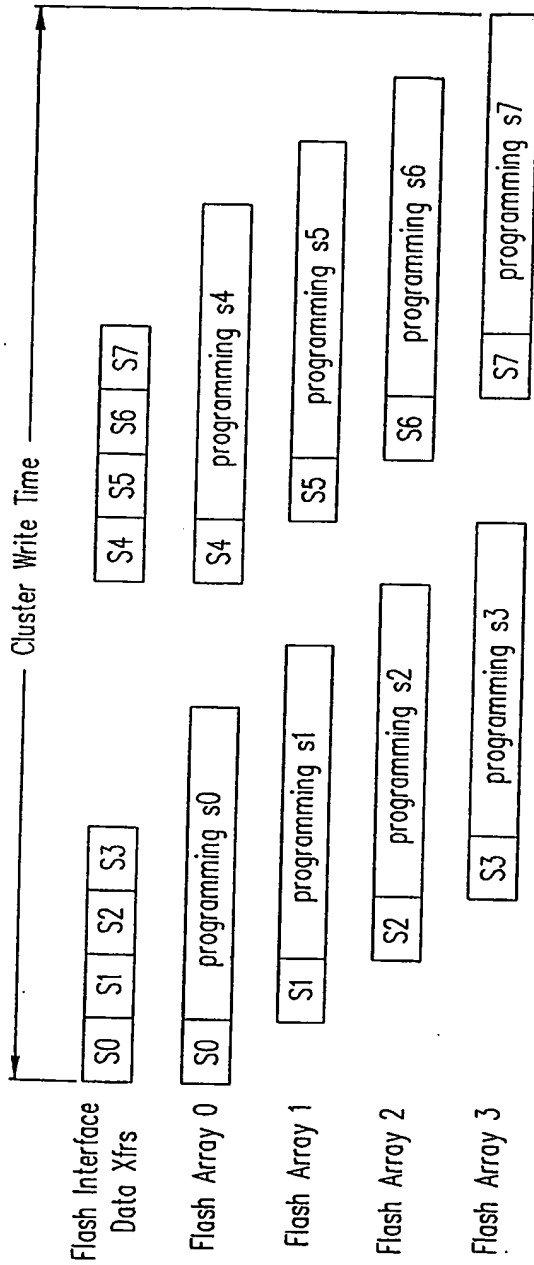


FIG. 9

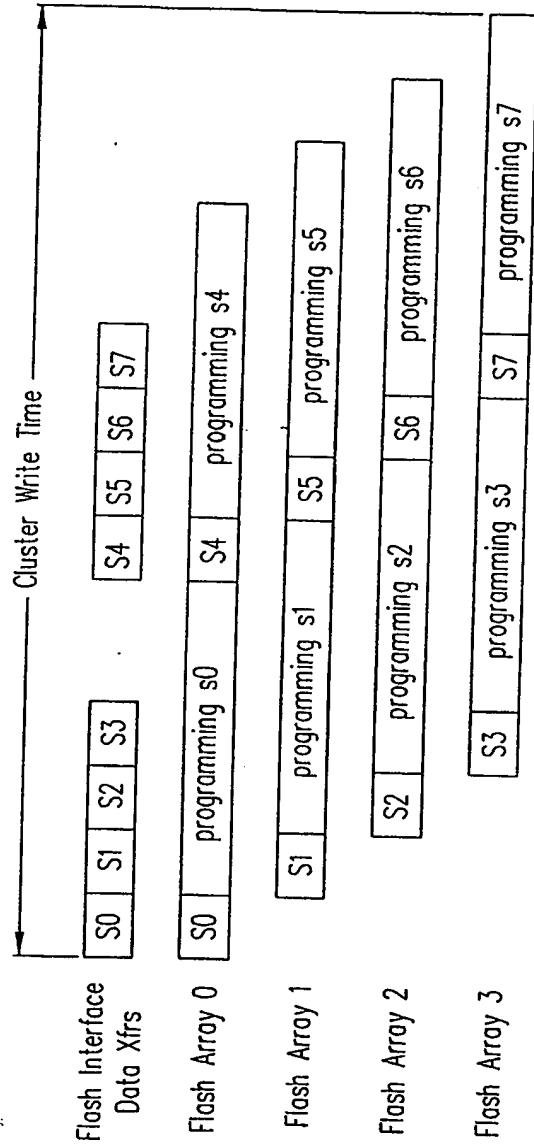


FIG. 10

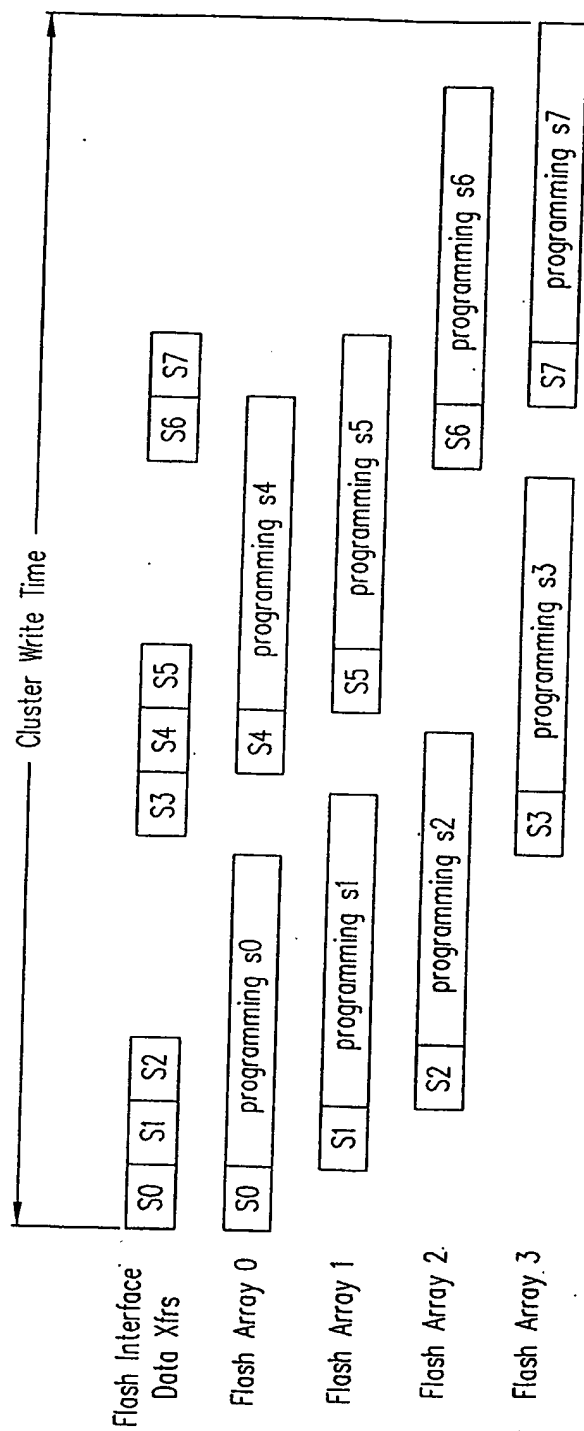


FIG. 11